Constructing Delaunay triangulations along space-filling curves

Citation for published version (APA):
Buchin, K. (2009). Constructing Delaunay triangulations along space-filling curves. In A. Fiat, & P. Sanders (Eds.), *Algorithms - ESA 2009 (17th Annual European Symposium, Copenhagen, Denmark, September 7-9, 2009. Proceedings)* (pp. 119-130). (Lecture Notes in Computer Science; Vol. 5757). Springer. https://doi.org/10.1007/978-3-642-04128-0_11

DOI:
10.1007/978-3-642-04128-0_11

Document status and date:
Published: 01/01/2009

Document Version:
Author's version before peer-review

Constructing Delaunay Triangulations along Space-Filling Curves

Kevin Buchin*

Abstract

Incremental construction con BRIO using a space-filling curve order for insertion is a popular algorithm for constructing Delaunay triangulations. So far, it has only been analyzed for the case that a worst-case optimal point location data structure is used which is often avoided in implementations. In this paper, we analyze its running time for the more typical case that points are located by walking. We show that in the worst-case the algorithm needs quadratic time, but that this can only happen in degenerate cases. We show that the algorithm runs in $O(n \log n)$ time under realistic assumptions. Furthermore, we show that it runs in expected linear time for many random point distributions.

1 Introduction

Delaunay triangulations (DTs) and their dual Voronoi diagrams are frequently used in many application areas, such as surface reconstruction, molecular modeling, and geographical information systems. They have been extensively studied in computational geometry and many different construction algorithms have been devised. Since its introduction in 2003 *Incremental Construction con BRIO* (biased randomized insertion order) [1] has been one of the favorite algorithms for constructing DTs. Points are inserted in rounds of increasing size which avoids full randomization. In a round the insertion order can be chosen, for which mostly space-filling curve (SFC) orders are used (see Fig. 1(d) for such an order). Already considered in the original article [1] (see also [20]), these orders have been have been popularized by Liu and Snoeyink [15] who used them in their program for constructing DTs of finite-precision input points. A variant of the algorithm is available as package in the *Computational Geometry Algorithms Library*[1] (CGAL) [6].

In the incremental construction, to insert a point it first has to be located in the current triangulation. When inserting points along a SFC order, this is typically done without an additional point location data structure. Using the spatial coherence of the order, a new point is located by *walking* from the previous point in the order, i.e., by traversing the triangulation data structure starting at this point. Incremental construction con BRIO with SFC orders have been tested thoroughly, and their running time on surface, protein, terrain and random data is linear or near-linear in experiments [1, 4, 15, 20]. The algorithm can be made asymptotically optimal by using a point location data structure like the conflict graph, but this not only requires additional space but also does not make use of the spatial coherence of the order. Without such a data structure non-trivial bounds on the running time were not known so far.

Most commonly the running time of DT algorithms is analyzed with respect to the worst-case point distribution. The drawback of such an analysis is that worst-case point sets might be degenerate and that the worst-case bound might not represent the running time on typical points

*Department of Mathematics and Computer Science, TU Eindhoven; *kbuchin@win.tue.nl*

[1] *http://www.cgal.org/*

well. On the other extreme, some DT algorithms have been analyzed with respect to the average-case running time on points drawn independently and uniformly from a unit square, or in higher dimensions from the unit d-cube [3, 8, 9, 10, 13, 16, 19]. This is an insightful alternative to the worst-case analysis, although such an input might be rather unlikely. Such an analysis can be strengthened by extending it to further random point distributions. However, except for trivial extensions to nearly uniform points, this has not been done for DT algorithms.

An alternative to the traditional worst- and average-case analysis are *realistic input models*. A global parameter for point sets which can often be bounded (in the size of the set) is its *spread*, i.e., the quotient between the largest and the smallest point to point distance. Frequently bounds on the spread result from minimum separation distances between the points and limited precision. In many cases the spread can be assumed to be polynomially bounded in the number of points. A further reason why the spread can be expected to be bounded, in particular when the points come from measurements, is noise in the data. *Smoothed analysis* [5, 18] models this by allowing arbitrary input point sets, but by performing an average-case analysis with the points perturbed by random noise. In the case of surface reconstruction a realistic assumption is that the surface is *well-sampled*, i.e., every surface point has one (but not too many) sample point close to it.

So far, realistic input models have to the best of our knowledge not been used explicitly for analyzing DT algorithms. They have been used to bound the complexity of DTs in $\mathbb{R}^3$. Although most three-dimensional point sets occurring seem to have DTs of linear size, their worst-case complexity is quadratic. Point sets in R^3 with spread Φ have complexity $O(\Phi^3)$ [11]. There are many results on well-sampled surfaces, for instance DTs of (ε, κ)-sampled polyhedra have linear complexity [2]. Recall that a point set is an (ε, κ)-sample of a polyhedron if all points on a facet of the polyhedron have a sample point on the same facet closer than ε to it, but not more than κ sample points on the same facet closer than 2ε to it.

Our results. Our aim is to give a theoretical explanation for the linear or near-linear running time in experiments of incremental construction con BRIO with SFC orders. As we will show, the worst-case running time is quadratic. We therefore turn to realistic and probabilistic models. We prove that the running time is $O(N \log \Phi)$ for an N-point set in the plane with spread Φ. This bound is tight for worst-case point sets as long as the spread is at least $w(\sqrt{N})$ (and at most exponential). Thus, if the spread is polynomially bounded then the running time is in $O(N \log N)$. This directly implies a similar bound for the smoothed complexity and can be easily extended to a bound for well-sampled surfaces. The bound also holds for points drawn from any typical random distribution, but for this case we even show a stronger bound. For independent identically distributed points from a large class of distributions a variant of the algorithm runs in linear expected time after computing a SFC order (which can be computed in linear expected time in a suitable model of computation). We give the explicit analysis for uniformly and for normally distributed points. Our results extend to higher dimensions, in which case the running time depends on the structural change of the incremental construction.

This is the first analysis of a DT algorithm for realistic input models. It is also the first probabilistic analysis of a DT algorithm that goes beyond uniformly distributed points. Besides the analysis of incremental search [10] it is the only probabilistic analysis for DT of points in higher dimensions. Proving linear expected running time for this algorithm also solves an open problem posed in the original con BRIO paper [1]. It is especially surprising that the algorithm achieves these running times without a point location data structure. So far, the fastest incremental construction algorithm without point location data structure was the *jump & walk* algorithm [9, 16] which runs in $O(n^{3/2})$ expected time for uniformly distributed points in a square (and close to

Algorithm 1: Incremental Construction along Space-Filling Curves

```
Input: Point set in $\mathbb{R}^d$
Output: Delaunay triangulation of the point set
1 Compute BRIO with SFC in rounds:
  1.1 Sample points to rounds (using coin flips with sampling ratio 1/2),
  1.2 Order points in a round using a space-filling curve order,
      for every other round use reversed order.
2 Incrementally construct Delaunay triangulation using order from Step 1:
  In each step do
  2.1 Locate new point from the previously inserted point by walking,
  2.2 Update Delaunay triangulation.
```

$O(n^{4/3})$ expected time in a 3-cube), with no worst-case guarantee except for the straightforward $O(n^2)$ bound. In addition to analyzing incremental construction con BRIO with SFC orders, we also present a generalized analysis of incremental constructions con BRIO which also applies to settings other than DTs.

2 Algorithm

The algorithm combines a *biased randomized insertion order (BRIO)* [1] and *space-filling curve (SFC) orders* (see Algorithm 1). In a BRIO points are grouped into rounds. Each point is independently assigned to the last round with probability 1/2. Points not assigned to the last round are assigned to the next to last round with the probability of 1/2 and so on [1]. After a logarithmic number of rounds an expected constant number of points remain, and we can therefore stop the sampling and assign the remaining points to the first round. If $p \in S_i$ denotes that the point p is inserted in round i or before $(i \geq 1)$ then the assignment can be described in terms of probabilities as $\mathrm{P}\left[p \in S_i \,|\, p \in S_{i+1}\right] = \frac{1}{2}$ for $1 \leq i < \lceil \log_2 N \rceil + 1$ and $\mathrm{P}\left[p \in S_{\leq \lceil \log_2 N \rceil + 1}\right] = 1$. For the analysis of the algorithm we will use the fact that the expected structural change, i.e., the total number of simplices created and deleted, using a BRIO is asymptotically bounded by the expected structural change using a randomized order (see Section 3.1).

Within a round we sort points along a space-filling curve. A SFC maps a 1-dimensional space onto a higher-dimensional space, e.g., the unit interval onto the unit square. We will use SFCs in the form of the SFC heuristic for the Euclidean traveling salesperson problem [17]. We demonstrate the SFC heuristic for this task by the example of the two-dimensional Hilbert curve [12]. Consider the following construction of a map from the unit interval to the unit square: Divide the unit interval into four intervals, divide the unit square into four squares, and assign each interval to one of the squares (see Fig. 1(a)). This process can be continued recursively and furthermore it can be done in such a way that neighboring intervals are assigned to neighboring squares. The first three steps of this construction are shown in Fig. 1(a-c). In the limit this yields a surjective, continuous map from the unit interval to the unit square. By its recursive construction the Hilbert curve maps an interval to a region with an area equal to the length of the interval. This is a property shared by many space-filling curves referred to as *bi-measure-preserving property*. Another property shared by many space-filling curves including the Hilbert curve is that they are *Hölder*-$1/d$ *continuous*. This means for a space-filling curve $\psi\colon [0,1] \to [0,1]^d$ that there is a constant C such that $\|\psi(s) - \psi(t)\| \leq C|s-t|^{1/d}$ for all $s, t \in [0,1]$.

For our purposes it suffices to repeat the subdivision process until there is only one point per square of the subdivision. In this example, for one of the squares one more subdivision step is necessary. Fig. 1(d) shows the resulting order. We call this order of the points a *space-filling curve order*. We will call the graph obtained by connecting the points in this order *space-filling curve tour*. For efficiently computing the SFC order of a point set in $\mathbb{R}^d$, we do $\log_{2^d} N$ subdivision steps at once. This results in $\Theta(N)$ cells in the subdivision. The order of the cells and the orientation of the curve in a cell can be stored in a look-up table. Using the floor function restricted to $\log N$ bits and radix sort, we can compute a SFC order of a point set with polynomially bounded spread in linear time since a constant number of rounds are sufficient in this case. In particular this yields a linear (expected) time[2] construction for the point sets we will consider. After computing the insertion order, we incrementally construct the Delaunay triangulation (DT) using the order. A point is located by a *straight line walk* from the previously inserted point, i.e, we trace the line segment from the previous point to the new point in the DT data structure.

Since we are mostly interested in the point location cost of the DT construction, we will in the following assume that the points are already given in a SFC order. In our experiments (see [4, Section 4.5]) the computation of the SFC order made up about 10% of the running time in two dimensions and less in three dimensions. The experiments also confirm that a sampling ratio smaller than 1/2 (in two dimensions between 1/10 and 1/4) speeds up the algorithm as has been already observed in earlier experiments [15, 20]. Our analysis easily generalizes to other sampling ratios.

3 General Analysis

3.1 Incremental Construction con BRIO revisited

Amenta, Choi, and Rote [1] introduced biased randomized insertion orders in the context of Delaunay triangulations of points sampled from a surface in $\mathbb{R}^3$. They consider point sets for which the expected complexity of the DT of a random sample of the point set is linear in the size of the sample. They prove for this case that the expected total update and point location cost with a history are for BRIOs asymptotically the same as for random orders. For our analysis we need to generalize their result to points in any dimension. We simplify their analysis by directly linking the costs for a construction with biased randomized insertion order to the costs for a randomized construction.

For a d-simplex with vertices in P and with s conflicting points in P let $p_B(s)$ and $p_R(s)$ denote the probabilities that the simplex occurs in an incremental construction with biased randomized insertion order and with randomized insertion order, respectively. For simplicity we assume that the

[2] Without the restricted floor function we get an additional factor $\log N$.

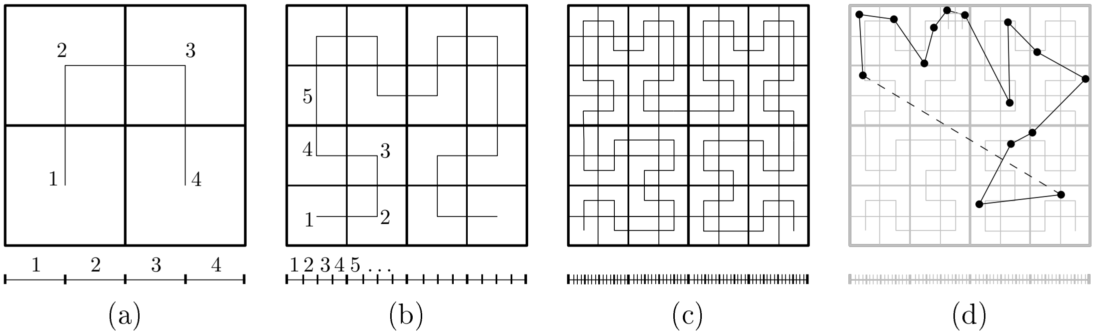


Figure 1: Hilbert curve and order.

sampling to rounds is not stopped after $\log(n)$ steps, but when no points remain (see [4, Proposition 3.7] for an analysis without this assumption). We bound $p_B(s)$ in terms of $p_R(s)$. This directly yields bounds for the costs determining the expected run-time of the construction, i.e., the expected *structural change* $\sum_{k=0}^{n} k_s p_B(s)$ and the expected *conflict change* $\sum_{k=0}^{n} s k_s p_B(s)$, where k_s denotes the total number of d-simplices with s conflicts. Note that the following lemma directly generalizes to arbitrary degree bounded configuration spaces and sampling ratios $1/\alpha$ by replacing (in the lemma and its proof) $d+1$ by the degree bound and 2 by α.

Lemma 3.1 *For a point set in* $\mathbb{R}^d$ *it holds that* $p_B(s) \leq 2^{d+1} p_R(s)$. ⊠[3]

3.2 Counting Intersections

In the following we develop a general scheme to count the number of intersections of a space-filling curve tour with a possibly changing Delaunay triangulation. Viewing this number as a double sum over the simplices of the DT and the line segments of the tour, there are two natural ways to count the intersections. In this section we will count for each simplex the number of line segments it intersects. More specifically, we will bound for each vertex of the DT the number of line segments of the SFC tour that might be intersected by a simplex with this vertex as a corner. This analysis allows us to focus on the structure of the tour. Alternatively, we could count for each line segment of the tour the number of simplices it intersects, which shifts the focus of the analysis to the structure of the DT. We will follow this alternative approach in Section 5.2.

Setup. Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be points in $\mathbb{R}^d$. Assume that we want to insert $y_1, \ldots, y_m$ into the Delaunay triangulation $\mathrm{DT}(x_1, \ldots, x_n)$ of the points $x_1, \ldots, x_n$. We insert $y_1, \ldots, y_m$ along a space-filling curve tour denoted by $\mathrm{T}(y_1, \ldots, y_m)$ which is given by a permutation $\pi\colon \{1, \ldots, m\} \to \{1, \ldots, m\}$. Let $f(x, \mathrm{DT})$ denote the number of d-dimensional faces incident to x in the Delaunay triangulation DT, e.g., in the plane the number of triangles incident to x. Let $F(\mathrm{DT})$ denote the total number of d-dimensional faces of DT and $C(\mathrm{DT}, \mathrm{T})$ the structural change when inserting the points of the tour T into DT in the order given by the tour. Let B_{y_i,y_j} denote the ball with the line segment (y_i, y_j) as a diameter. Furthermore, let $b(x, \mathrm{T}(y_1, \ldots, y_m)) := \sum_{i=1}^{m-1} \mathbf{1}_{B_{y_i,y_{i+1}}}(x)$, i.e., the number of balls around tour segments in which x lies. In a probabilistic setting we denote the random variables corresponding to $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ as $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, respectively.

Counting Scheme. For points in general position the faces of the Delaunay triangulation intersected by tour segments are $(d-1)$-dimensional or d-dimensional with these two cases alternating along the tour segment. Of these, we will count the d-dimensional faces.

Let I be the number of intersections between d-simplices of the current Delaunay triangulation and line segments of the space-filling curve tour. We will consider two scenarios: In the first, we directly insert a new point after we located it. This corresponds to the situation in Algorithm 1. For the line segment $y_{\pi(i)} y_{\pi(i+1)}$ we count the number of intersections with $\mathrm{DT}(x_1, \ldots, x_n, y_{\pi(1)}, \ldots y_{\pi(i)})$ $(1 \leq i < m)$. In a second scenario, we will simply count the number of intersections between $\mathrm{DT}(x_1, \ldots, x_n)$ and $\mathrm{T}(y_1, \ldots, y_m)$. Most of the analysis will handle both scenarios simultaneously.

We split the number of intersections into $I = I_1 + I_2$ where

- I_1 is the number of intersections where the d-simplex is in conflict with one of the endpoints of the tour segment,

[3] We use ⊠ to indicate that the proof has been postponed to Appendix A.

- I_2 is the number of intersections where the d-simplex is not in conflict with the endpoints of the tour segment.

When we want to solely refer to the second scenario, we will use $\widehat{I_2}$ to denote the second number instead of I_2.

Lemma 3.2 *Let Δ be a d-simplex and s a line segment intersecting Δ. If the endpoints of s lie outside of the circumsphere of Δ then the ball with s as diameter contains a vertex of Δ.* ⊠

Consider a fixed line segment $(y_{\pi(i)}, y_{\pi(i+1)})$ on the space-filling curve tour. By Lemma 3.2 any d-face of the DT intersecting this segment and not in conflict with one of the endpoints of the tour segment must have one vertex in the ball with the tour segment as diameter. Thus, for any intersection counted in I_2 the corresponding Delaunay simplex has a vertex in the ball with the corresponding tour segment as a diameter. We will use this to bound I_2.

Bounding I_1. A Delaunay face in conflict with a vertex of the tour needs to be counted at most once for each tour segment adjacent to the vertex, i.e., at most twice for the vertex. In the first scenario it is actually only counted once, since it is no longer in the triangulation after the insertion of the vertex. We can bound the cost induced by these faces by the structural change, i.e., $I_1 \in O(C(\mathrm{DT}(x_1, \ldots, x_n), \mathrm{T}(y_1, \ldots, y_m)))$.

Bounding I_2. We bound I_2 by counting for each vertex of the Delaunay triangulation in a ball of a tour segment the total number of d-simplices at this vertex. In the first scenario, i.e., if we insert points while traversing the tour, we have

$$I_2 \leq \sum_{i=1}^{m-1} \sum_{j=1}^{n} \mathbf{1}_{B_{y_{\pi(i)}, y_{\pi(i+1)}}}(x_j) f(x_j, DT(x_1, \ldots, x_n, y_{\pi(1)}, \ldots, y_{\pi(i)})) \\ + \sum_{i=1}^{m-1} \sum_{j=1}^{i-1} \mathbf{1}_{B_{y_{\pi(i)}, y_{\pi(i+1)}}}(y_{\pi(j)}) f(x_j, DT(x_1, \ldots, x_n, y_{\pi(1)}, \ldots, y_{\pi(i)})).$$

In the second scenario this bound is simply

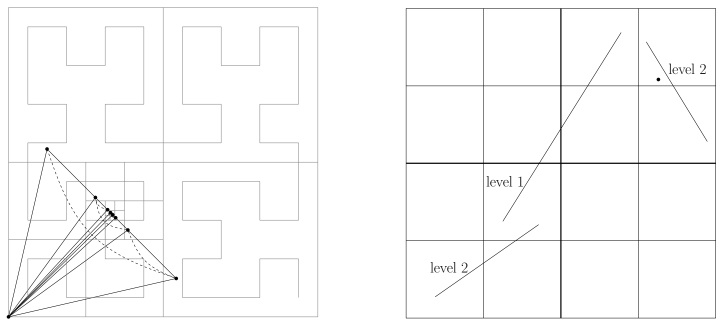
$$\widehat{I_2} \leq \sum_{i=1}^{m-1} \sum_{j=1}^{n} \mathbf{1}_{B_{y_{\pi(i)}, y_{\pi(i+1)}}}(x_j) f(x_j, DT(x_1, \ldots, x_n)) = \sum_{j=1}^{n} b(x_j, T(y_1, \ldots, y_m)) f(x_j, DT(x_1, \ldots, x_n)).$$

Proposition 3.3 *Let $P := \{x_1, \ldots, x_n, y_1, \ldots, y_m\}$ and $b_m := \max_{z \in P} b(z, T(y_1, \ldots, y_m))$. Then*

$$I_2 \leq b_m(d+1)(F(DT(x_1, \ldots, x_n)) + C(DT(x_1, \ldots, x_n), T(y_1, \ldots, y_m))).$$

Proof: Any vertex is covered by at most b_m balls of the tour. Counting a simplex b_m times for each incident vertex counts it $b_m(d+1)$ times. Thus, we can bound I_2 by $b_m(d+1)$ times the total number of simplices occurring. The number of simplices is bounded by $F(DT(x_1, \ldots, x_n)) + C(DT(x_1, \ldots, x_n), T(y_1, \ldots, y_m))$. ■

Proposition 3.3 gives a worst-case bound on I_2. The straightforward generalization of the proposition to a probabilistic setting, would replace b_m by the *expected maximum coverage*. In the following we show that if we turn to the second scenario, i.e., do not insert points directly, we can replace b_m by the typically smaller *maximum expected coverage* instead.



(a) Points with $\Omega(n^2)$ running time.

(b) Levels of tour segments

Figure 2: Lower and upper bound constructions.

Proposition 3.4 *Let $X_1, \ldots, X_n, Y_1, \ldots, Y_m \in D \subseteq \mathbb{R}^d$ be independent random variables.. Let*

$$\widehat{b_m} := \sup_{x \in D} E\left[b(x, T(Y_1, \ldots, Y_m))\right] \text{ and } F_n := E\left[F(DT(X_1, \ldots, X_n))\right].$$

Then $E\left[\widehat{I_2}\right] \leq (d+1)\widehat{b_m}\, F_n$. ☒

4 Analysis for Bounded Spread

4.1 Lower Bound

In this section we will focus on DTs in the plane. In two dimensions the running time of Algorithm 1 is trivially in $O(n^2)$, since the time needed to locate one point is at most linear. Unfortunately, for worst-case point sets the algorithm indeed needs quadratic time, as we show next. We construct a point set for the Hilbert curve. Fig. 2(a) shows the point set for $N = 9$. The first point is placed at $(0, 0)$. All further points are placed on the line $y = 2/3 - x$. Note that by adding a small offset to the points, they can be placed in strictly convex position instead. The x-coordinates of these $N - 1 = 2K$ points are $1/8, 1/4 + 1/(4 \cdot 8), \ldots, \sum_{i=1}^{K-1} 1/4^i + 1/(4^{K-1} \cdot 8)$ and $2/3 - 1/8, 2/3 - (1/4 + 1/(4 \cdot 8)), \ldots, 2/3 - (\sum_{i=1}^{K-1} 1/4^i + 1/(4^{K-1} \cdot 8))$. The points are chosen such that the space-filling curve tour first traverses the points closest to the diagonal $(0, 0), (1, 1)$, going outward from there. This can be seen from the self-similar structure of the point set, i.e., the situation in a sub-square is essentially the same as in the original square (with two points less). Now we pair up the points on the line by their distance to the diagonal. Any such pair has probability $1/4$ to be inserted in the last round, and the ith pair intersects $2(i - 1)$ lines. Thus the expected number of intersections is $\Omega(n^2)$ which dominates the running time.

Our worst-case example is highly degenerate. Most notably it has exponential spread. We therefore study how the running time parameterizes in terms of the spread. Adapting the worst-case above yields the following bound.

Proposition 4.1 *For $\Phi(N) \in \omega(\sqrt{N}) \cap 2^{O(N)}$ there are point sets of size N for which the spread is at most $\Phi(N)$ and the running time of Algorithm 1 is in $\Omega(N \log \Phi(N))$.*

Proof: We place instances of the construction above with $k = \log(\Phi(N)\sqrt{N})$ points on a $\sqrt{N/k} \times \sqrt{N/k}$ grid. The total number of intersections occuring in the last round is in $\Omega(Nk) = \Omega(N \log \Phi(N))$, since $\Phi(N) \in \omega(\sqrt{N})$. ■

4.2 Upper Bound

In the following we show an upper bound matching the lower bound.

Proposition 4.2 *Algorithm 1 runs in $O(|P| \log \Phi(P))$ time in the plane.*

Proof: We assign a *level* to each edge of $\mathrm{T}(y_1, \ldots, y_m)$ according to the highest subdivision level (counting from coarse to fine) for which the edge is still contained in a single cell (see Fig. 2(b) for an example). Any point can be in the ball of at most a constant number of edges per level, for instance the point in the upper right of Fig. 2(b) cannot be in a ball corresponding to a level-2-edge with vertices in the lower left square. Further, the number of levels is in $O(\log \Phi(N))$, which yields the claimed running time using Proposition 3.3. ■

Extensions. Many point sets have polynomially bounded spread, in which case our bound implies that the algorithm runs in $O(N \log N)$ time. In higher dimensions the complexity of the DT is not necessarily linear, so we get as bound on the running time $O(C(P) \log \Phi(P))$, where $C(P)$ denotes the structural change. In a smoothed analysis the noise added (as long as it is not exponentially small) will bound the expected smallest point-to-point distance and therefore the expected spread, if the largest point-to-point distance is bounded. Thus, we again obtain a $O(N \log N)$ running time. For well-sampled domains we can typically restrict the number of levels we need to consider. For instance for (ε, κ)-sampled polyhedra we can subdivide until cells are covered by a 2ε-ball. We then know that only κ points per cell remain. Thus, we get a running time of order $O(n(\log(1/\varepsilon) + \kappa))$.

5 Average-Case Analysis

5.1 Structure of Random Space-Filling Curves

In the previous section we obtained a running time of $O(N \log N)$ for typical inputs. To prove even stronger bounds we turn to an average-case analysis. For this we will use Proposition 3.4, which only holds if we do not insert points directly. We therefore consider the following variant of Algorithm 1: The point location of a round is done in two steps. First, points are located in the DT of the points of the previous rounds by a walk along the SFC order. Second, points are located from the location found by the walk using the history. Note that for this we only need to maintain the history of the current round. We have the choice of inserting the points in a random order or in the order given by the space-filling curve. In the first case we directly obtain an expected constant point location cost [7]. In the second case the same follows from Remark B.1 in the appendix.

Let B_T be a ball chosen uniformly at random from the balls along the tour with m vertices. To prove $\mathrm{E}\left[\widehat{I_2}\right] \in O(n)$ it suffices to prove that for all $x \in D$ it holds that $\mathrm{P}\left[x \in B_T\right] \in O(1/m)$, where D is the domain from which the X_i are drawn $(1 \leq i \leq n)$. Now, $\mathrm{P}\left[x \in B_T\right]$ does not depend on properties of the Delaunay triangulation, thus we have reduced the problem to a problem on properties of the tour. To bound $\mathrm{P}\left[x \in B_T\right]$ we will now use the Hölder-continuity of space-filling curves. Therefore, the following analysis will only hold for space-filling curve orders.

The following argument suggests that $\mathrm{P}\left[x \in B_T\right] \in O(1/m)$ for space-filling curve orders if the point distribution considered is "sufficiently smooth": if $\psi : [0,1] \to [0,1]^d$ is the space-filling curve used then the probability that the preimage $\psi^*(x)$ of a point x falls into the interval $[\psi^*(Y_{\pi(i)}), \psi^*(Y_{\pi(i+1)})]$ between two points defining a ball is $1/(m-1)$. Thus the probability that a point lies in the image of the interval under ψ is again $1/(m-1)$ and this image covers a similar area as the ball. This notion of "similar area" is difficult to capture and we will argue directly using the probability distribution over the space from which the points are drawn.

We want to bound $\mathrm{P}\left[x \in B_T\right]$. First it is important to consider how the space-filling curve was computed. If the points come from a certain region we can simply compute the space-filling curve based on a subdivision of this region. But for points from an unbounded region, like in the case of the normal distribution, the bounding cube for the space-filling curve depends on the actual points. For simplicity we will assume that the bounding cube is chosen as $[-u, u]^d$ where u is the largest occurring coordinate, i.e., the largest L_∞-norm of a point. For a space-filling curve $\psi \colon [0, 1] \to [0, 1]^d$ we denote by $\hat{\psi} \colon [0, 1] \to [-u, u]^d$ the scaled space-filling curve. The mapping $\hat{\psi}$ is Hölder continuous with exponent $1/d$ and Hölder constant $c_{\hat{\psi}} = 2u \cdot c_\psi$, i.e., for $t_1, t_2 \in [0, 1]$ we have $\left\|\hat{\psi}(t_1) - \hat{\psi}(t_2)\right\| \leq c_{\hat{\psi}} \left\|t_1 - t_2\right\|^{1/d}$. We denote by $\hat{\psi}^* \colon [-u, u]^d \to [0, 1]$ the selection of preimages according to ψ^*. The following lemma provides a bound on the length of a tour edge in this setting.

Lemma 5.1 *Let $Y_1, \ldots, Y_m$ be independent identically distributed random variables in $\mathbb{R}^d$ with Lebesgue density function g_{Y_1}. Let $\psi : [0, 1] \to [0, 1]^d$ be a Hölder continuous and bi-measure preserving space-filling curve with Hölder constant c_ψ. Let L be a random tour segment of a space-filling curve tour through $Y_1, \ldots, Y_m$ based on $\hat{\psi}$. Then for all $\ell > 0$*

$$P\left[|L| > \ell\right] \leq \int_{\mathbb{R}^d} g_{Y_1}(y) \left(1 - \frac{d}{c_{\psi}^d} \int_{[0,\ell]} s^{d-1} \min\left\{g_{Y_1}(y') \mid \left\|y - y'\right\| < s\right\} ds\right)^{m-1} d\lambda^d(y).$$

⊠

Using Lemma 5.1 we can bound $\mathrm{P}\left[x \in B_T\right]$ by

$$\int_{\mathbb{R}^d} g_{Y_1}(y) \left(1 - \frac{d}{c_{\psi}^d} \int_{[0,\|x-y\|]} s^{d-1} \min\left\{g_{Y_1}(y') \mid \left\|y - y'\right\| < s\right\} ds\right)^{m-1} d\lambda^d(y).$$

Two examples for distributions handled by Lemma 5.1 are uniformly distributed points in $[0, 1]^d$ and normally distributed points in the plane. For normally distributed points we apply the lemma to all points except a few (an expected logarithmic number of points) far away from the center of the distribution. To handle the remaining points we assume for the sake of the analysis that an additional point location structure like Kirkpatrick's point location hierarchy [14] is used.

Theorem 5.2 *The* incremental construction along space-filling curves *(using the history of a round) computes the Delaunay triangulation of points drawn independently and uniformly from a d-cube in linear expected time.* ⊠

Theorem 5.3 *The* incremental construction along space-filling curves *(using the history of a round and an $O(\log n)$ point location data structure) computes the Delaunay triangulation of points drawn independent identically normally distributed in the plane in linear expected time.* ⊠

5.2 Structure of Random Delaunay Triangulations

We next give an alternative average-case analysis. It focuses on the structure of the DT rather than the structure of the SFC. We analyze Algorithm 1 for points drawn independently and uniformly from a bounded convex region (of area 1) in the plane. To analyze the run-time it suffices to analyze the run-time of the last round. This bound can then be applied to the other rounds with the number of points being a random variable depending on the round. We assume that at the beginning of the last round n points have already been inserted into the DT, while the m points from the last round are to be inserted. The points are located by traversing the Delaunay triangulation along a space-filling curve order. Therefore, the time for locating the points is proportional to the number of intersections between the order and the Delaunay triangulation. Let L be a line segment that is not too close (no closer than $c\sqrt{\log n/n}$ for a suitable constant c) to the boundary of the bounded convex region and that is independent from the (independently and uniformly distributed) points of the DT. Then the expected number of intersections between L and the DT is in $O(1+\sqrt{n}|L|)$ [9]. Now, a Hölder-1/2 continuous space-filling curve order of m points in a bounded region in the plane yields a walk through the points of length $O(\sqrt{m})$ [17]. Combining these two results gives an expected number of intersections in $O(m+\sqrt{nm})$, but there are two pieces missing in this argument. First, points close to the boundary are not handled. Second, points are inserted during the walk. Therefore the DT changes and depends on the points to be inserted and their insertion order. In this extended abstract we only handle the first problem, i.e., points close to the boundary and refer to [4, Section 3.4.3] for the second.

We handle points close to the boundary by showing that edges in the DT are either short or near to the boundary. For this we use the following lemma, which generalizes a corresponding lemma for the case of points in the unit square [8]. See [4, Lemma 3.17] for a proof of the generalized lemma.

Lemma 5.4 *Let $n > 2$ points be distributed independently and uniformly in a bounded convex region C in the plane. Denote by $D_{w,\ell}$ the event that the Delaunay triangulation of the points contains an edge which has a point on it with distance at least w to the boundary of C and which is longer than ℓ. For any $t > 1$ and $\ell \geq tw$ $P[D_{w,\ell}] \leq n^2 e^{-(n-2)w\ell\sqrt{1-1/t^2}/2}$. In particular, if $\ell \geq 3w$ and $w\ell \geq 6\sqrt{2}\log n/(n-2)$ then $P[D_{w,\ell}] \leq 1/n^2$.*

This gives us a bound on the number of Delaunay edges that can intersect the line segments of the tour:

Lemma 5.5 *The expected number of intersections of a Delaunay triangulation of n points and a tour along a Hölder-$1/2$ space-filling curve through m points where the $m+n$ points are distributed independently and uniformly in a bounded convex region C is in $O(m+\sqrt{mn})$ and therefore linear in the total number of points.*

Proof: For line segments of the tour not near the boundary we already have an $O(m+\sqrt{mn})$ bound on the number of intersections with the Delaunay triangulation. We therefore only need to consider the line segments of the tour near the boundary. Without loss of generality, the area of C is 1.

The expected number of points near the boundary is bounded by $m' := c_0\,|\partial C|\,m\,\sqrt{\log n/n}$. Each of these points can be adjacent to two line segments resulting in an upper bound of $k := 2m'$ on the expected number of line segments with at least one point near the boundary. Each of these line segments can contribute one further endpoint resulting in a bound of $3m'$ on the expected number of endpoints. These line segments are also part of a space-filling curve tour through only these endpoints. Therefore, by the $O(\sqrt{m})$ bound on the length of the tour through the points ordered along the space-filling curve and Jensen's inequality the expected total length of these line segments is bounded by $\ell_\Sigma := c_T\sqrt{3m'}$ for a constant c_T depending on the space-filling curve.

For $n > 2$ we choose $w := (\log n/(n-2))^{2/3}$ and $\ell := 6\sqrt{2}(\log n/(n-2))^{1/3}$. Then for $n > 2$, $\ell \geq 3w$ and $w\ell \geq 6\sqrt{2}\log n/(n-2)$. Thus, we have by Lemma 5.4, that with probability at least $1-\frac{1}{n^2}$ line segments are closer than w to the boundary or shorter than l. If this event occurs, only Delaunay edges with endpoints which have distance at most ℓ to one of the line segments of the tour, or distance at most w from the boundary can intersect the space-filling curve tour. If this event does not occur, then in the worst case a tour segment is intersected by all edges of the triangulation, which gives $O(1/n)$ intersections in expectation. In total, this contributes at most $O(m/n)$ expected intersections.

In the case that the event occurs, for a single line segment L the region, in which both endpoints of an intersecting Delaunay edge must lie, is the union of the ℓ-neighborhood of L and the w-neighborhood of the boundary of C. The area of this region is bounded from above by $|\partial C|w + \pi\ell^2 + 2\ell|L|$. Therefore the expected number of endpoints of edges that intersect a line segment L is bounded by $n(|\partial C|w + \pi\ell^2 + 2\ell|L|)$. Consider the subgraph of the Delaunay triangulation induced by this vertex set. Because of planarity there are at most three times that many edges intersecting L as in this subgraph. For k line segments of total expected length L_Σ this yields a bound of $3n(|\partial C|kw + \pi\ell^2 k + 2\ell L_\Sigma)$ on the number of intersecting edges. Inserting the values for w, ℓ, k, and ℓ_Σ gives that the number of intersections near the boundary can be bounded by

$$\begin{aligned} 3n(|\partial C|2m'w + \pi\ell^2 2m' + 2\ell c_T\sqrt{3m'}) &\le 6nm'(\log n/n)^{2/3}(|\partial C| + \pi) + 18c_T\sqrt{6}nm'^{1/2}(\log n/n)^{1/3} \\ &\in O\left(m\frac{\log^{7/6} n}{n^{1/6}} + m^{1/2}n^{5/12}\log^{7/12} n\right) \subset O(m + \sqrt{mn}). \end{aligned}$$

■

As mentioned, Lemma 5.5 does not handle the fact that points are already inserted into the DT during the walk. Because of space constraints we omit the generalization of the lemma to this case. The generalized version bounds the expected number of intersections between the space-filling curve walk and the DT by $O(m^2/n + n)$ even if the DT is updated after walking steps (see [4, Lemma 3.26]). This yields the following theorem.

Theorem 5.6 *Using a biased randomized insertion order and, in each round, walking along a Hölder-*$1/2$*, bi-measure preserving space-filling curve, the incremental construction algorithm runs in linear expected time for points distributed independently and uniformly at random in a bounded convex region on a real-RAM with the floor function available and a word size of at least* $\log N$.

For a self-contained account we can use the non-generalized Lemma 5.5 to bound the expected run-time of a variant of Algorithm 1. In this variant the point location of a round is done again using the history of a round.

Acknowledgments. The author would like to thank Günter Rote and Scot Drysdale and Maike Buchin for many helpful ideas.

References

[1] N. Amenta, S. Choi, and G. Rote. Incremental constructions con BRIO. In *Proc. 19th Annu. ACM Sympos. Comput. Geom.*, pages 211–219. ACM Press, 2003.

[2] D. Attali and J.-D. Boissonnat. A linear bound on the complexity of the Delaunay triangulation of points on polyhedral surfaces. *Discrete Comput. Geom.*, 31(3):369–384, 2004.

[3] J. L. Bentley, B. W. Weide, and A. C. Yao. Optimal expected-time algorithms for closest-point problems. *ACM Trans. Math. Softw.*, 6:563–580, 1980.

[4] K. Buchin. *Organizing Point Sets: Space-Filling Curves, Delaunay Tessellations of Random Point Sets, and Flow Complexes.* PhD thesis, Free University Berlin, 2007. http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000003494.

[5] V. Damerow, F. Meyer auf der Heide, H. Räcke, C. Scheideler, and C. Sohler. Smoothed motion complexity. In *Proc. 11th Annu. European Sympos. Algorithms*, pages 161 – 171,, 2003.

[6] C. Delage. Spatial sorting. In CGAL Editorial Board, editor, *CGAL User and Reference Manual.* 2007.

[7] O. Devillers. Randomization yields simple o(n log* n) algorithms for difficult omega(n) problems. *Int. J. Comput. Geometry Appl.*, 2(1):97–111, 1992.

[8] L. Devroye, C. Lemaire, and J.-M. Moreau. Expected time analysis for Delaunay point location. *Comput. Geom. Theory Appl.*, 22:61–89, 2004.

[9] L. Devroye, E. Mücke, and B. Zhu. A note on point location in Delaunay triangulations of random points. *Algorithmica*, 22:477–482, 1998.

[10] R. A. Dwyer. Higher-dimensional Voronoi diagrams in linear expected time. *Discrete Comput. Geom.*, 6(4):343–367, 1991.

[11] J. Erickson. Dense point sets have sparse Delaunay triangulations: or "...but not too nasty". In *Proc. 13th Annu. ACM-SIAM Sympos. Discrete Algorithms*, pages 125–134, 2002.

[12] D. Hilbert. Ueber die stetige Abbildung einer Linie auf ein Flächenstück. *Math. Ann.*, 38:459–460, 1891.

[13] J. Katajainen and M. Koppinen. Constructing Delaunay triangulations by merging buckets in quadtree order. *Fundam. Inform.*, 11:275–288, 1988.

[14] D. G. Kirkpatrick. Optimal search in planar subdivisions. *SIAM J. Comput.*, 12(1):28–35, 1983.

[15] Y. Liu and J. Snoeyink. A comparison of five implementations of 3d Delaunay tesselation. In J. E. Goodman, J. Pach, and E. Welzl, editors, *Combinatorial and Computational Geometry*, volume 52 of *MSRI Publications*, pages 439–458. Cambridge University Press, 2005.

[16] E. P. Mücke, I. Saias, and B. Zhu. Fast randomized point location without preprocessing in two- and three-dimensional Delaunay triangulations. *Comput. Geom. Theory Appl.*, 12(1-2):63–83, 1999.

[17] L. K. Platzman and J. J. Bartholdi, III. Spacefilling curves and the planar travelling salesman problem. *J. ACM*, 36(4):719–737, 1989.

[18] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004.

[19] P. Su and R. Drysdale. A comparison of sequential Delaunay triangulation algorithms. *Comput. Geom. Theory Appl.*, 7:361–386, 1997.

[20] S. Zhou and C. B. Jones. HCPO: an efficient insertion order for incremental Delaunay triangulation. *Inf. Process. Lett.*, 93(1):37–42, 2005.

A Omitted Proofs

A.1 Proof of Lemma 3.1

Proof: For $s = 0$ we have $p_B(s) = 1 = p_R(s) \le 2^{d+1} p_R(s)$. In the following we consider a d-simplex Δ with $s > 0$ conflicts.

- $U_i := \{$All vertices of Δ appear in the ith to last round or before.$\}$,
- $V_i = \{$The first point in conflict with Δ appears in the ith to last round.$\}$,

where $i \in \mathbb{N}$, and $i = 1$ refers to the last round. For instance, V_2 is the event that all points in conflict with Δ are inserted in the last two rounds and at least one of them is inserted in the second to last round. The disjoint union $\sum_{i=1}^{\infty} V_i$ has probability 1. If V_i holds for an $i \in \mathbb{N}$ then Δ can only appear during the construction if also U_i holds. Now,

$$p_B(s) \;\le\; \mathrm{P}\left[\sum_{i=1}^{\infty}(V_i \cap U_i)\right] \;=\; \sum_{i=1}^{\infty} \mathrm{P}\left[V_i \cap U_i\right] \;=\; \sum_{i=1}^{\infty} \mathrm{P}\left[V_i\right] \mathrm{P}\left[U_i\right]$$

where the last equality holds by the independence of V_i and U_i for $i \in \mathbb{N}$. Next,

$$\mathrm{P}\left[U_{i+1}\right] \;=\; \mathrm{P}\left[U_{i+1} \cap U_i\right] \;=\; \mathrm{P}\left[U_{i+1} \,|\, U_i\right] \mathrm{P}\left[U_i\right] \;=\; \frac{1}{2^{d+1}} \mathrm{P}\left[U_i\right]$$

by the sampling condition. Therefore,

$$p_B(s) \;\le\; \sum_{i=1}^{\infty} 2^{d+1} \mathrm{P}\left[V_i\right] \mathrm{P}\left[U_{i+1}\right] \;=\; 2^{d+1} \mathrm{P}\left[\sum_{i=1}^{\infty}(V_i \cap U_{i+1})\right],$$

i.e., $p_B(s)$ is bounded by 2^{d+1} times the probability that all vertices of Δ are inserted in rounds strictly after the points conflicting with Δ. This probability does not depend on the order within rounds. Consider an incremental construction con BRIO with random order within rounds. This gives a randomized incremental construction. The event that all points conflicting with Δ appear in rounds before all vertices of Δ is included in the event that all conflicting points appear before all vertices. Thus, we get

$$\mathrm{P}\left[\sum_{i=1}^{\infty}(V_i \cap U_{i+1})\right] \le p_R(s) \tag{1}$$

and therefore $p_B(s) \;\le\; 2^{d+1} p_R(s)$. ■

A.2 Proof of Lemma 3.2

Proof: Let U be the circumsphere of Δ and let K be the ball with s as a diameter. Without loss of generality we can assume that the endpoints of s lie on U. Otherwise we can shrink s by which we also shrink K. We want to show that K contains a vertex of Δ. We can assume that the radius of K is less than the radius of U. Otherwise we have $U = \partial K$ and all vertices of Δ lie in K.

Now consider the spherical cap $U \cap K$. Its boundary lies in a hyperplane H and the cap itself in a half-space H^+ defined by H. In particular,

$$U \cap H^+ = (U \cap K) \cap H^+ \subset K.$$

Since s intersects Δ there is a vertex x of Δ in H^+. For this vertex we have

$$x \in U \cap H^+ \subset K.$$

■

A.3 Proof of Proposition 3.4

Proof: In the following we abbreviate $T(Y_1, \ldots, Y_m)$ by T and $DT(X_1, \ldots, X_n)$ by DT. We have

$$\begin{aligned} \mathrm{E}\left[I_2\right] &\le \mathrm{E}\left[\sum_{i=1}^{n} b(X_i, T) \cdot f(X_i, DT)\right] \\ &\le \sum_{i=1}^{n} \mathrm{E}\left[b(X_i, T) \cdot f(X_i, DT)\right] \\ &\le \sum_{i=1}^{n} \mathrm{E}\left[\mathrm{E}\left[b(X_i, T) \,|\, X_i\right] \cdot f(X_i, DT)\right]. \end{aligned}$$

Since X_i is independent of $Y_1, \ldots, Y_m$, we have $\mathrm{E}\left[b(X_i, T) \,|\, X_i = x_i\right] = \mathrm{E}\left[b(x_i, T)\right]$. Since for all x_i $\mathrm{E}\left[b(x_i, T)\right] \le \sup_{x \in D} \mathrm{E}\left[b(x, T(Y_1, \ldots, Y_m))\right] = b_m$ we have $\mathrm{E}\left[b(X_i, T) \,|\, X_i = x_i\right] \le b_m$. Thus,

$$\begin{aligned} \mathrm{E}\left[I_2\right] &\le b_m \sum_{i=1}^{n} \mathrm{E}\left[f(X_i, DT)\right] \\ &= b_m(d+1)\mathrm{E}\left[F(DT)\right] \\ &= (d+1) b_m F_n. \end{aligned}$$

■

A.4 Proof of Lemma 5.1

Proof:

Let Y (from $Y_1, \ldots, Y_m$) be the first point of the random segment L. To bound $\mathrm{P}\left[|L| > \ell\right]$ we first consider $\mathrm{P}\left[|L| > \ell \,|\, Y = y\right]$. For this we use the Hölder continuity and bi-measure preserving property of the space-filling curve. A tour segment of length ℓ yields that the empty region has a volume of order at least ℓ^d. In turn, an area of ℓ'^d between two points yields that the distance between the points is at most of order ℓ'.

In the case that $|L| > \ell'$ for an $\ell' > 0$ we get

$$\left\| \hat{\psi}^*(Y) - \hat{\psi}^*(Y') \right\| \ \ge \ \left(\frac{\ell'}{c_{\hat{\psi}}} \right)^d,$$

i.e., a tour segment longer than ℓ' implies an interval of length at least $(\ell'/c_{\hat{\psi}})^d$ in which no other preimage lies. By the bi-measure-preserving property of the space-filling curve we can consider the probabilities over [0,1]. Because of the scaling, a density of λ on the side of the points in $[-u, u]^d$ corresponds to a density of $(2u)^d \lambda = (c_{\hat{\psi}}/c_{\psi})^d \lambda$ on the side of the preimages in $[0, 1]$.

If the measure corresponding to an interval is a then the probability for one point to be outside of the interval is $1-a$ and for all $m-1$ remaining points is $(1-a)^{m-1}$.

Using the bi-measure-preserving property, now in the other direction, we can bound the change of the distribution for $t_0, t_0+t \in [0,1]$ by

$$g_{Y_1}(\hat{\psi}(t_0+t)) \geq \min\left\{ g_{Y_1}(\hat{\psi}(t')) \mid \left\| \hat{\psi}(t_0) - \hat{\psi}(t') \right\| \leq c_{\hat{\psi}} t^{1/d} \right\}.$$

This yields that

$$a \geq \int_{[0,(\ell/c_{\hat{\psi}} t)^d]} \left(\frac{c_{\hat{\psi}}}{c_{\psi}} \right)^d \min\left\{ g_{Y_1}(y') \mid \|y'-y\| \leq c_{\hat{\psi}} t^{\frac{1}{d}} \right\} d\lambda(t).$$

Substituting the integration variable t by $s = c_{\hat{\psi}} \cdot t^{1/d}$ we get

$$a \geq \frac{d}{c_{\psi}^d} \int_{[0,\ell]} s^{d-1} \min\left\{ g_{Y_1}(y') \mid \|y'-y\| \leq s \right\} d\lambda(s).$$

Thus we have a bound on a not depending on u. Now

$$\mathrm{P}\left[|L| > \ell \mid Y = y\right] \leq \left(1 - \frac{d}{c_{\psi}^d} \int_{[0,\ell]} s^{d-1} \min\left\{ g_{Y_1}(y') \mid \|y'-y\| \leq s \right\} d\lambda(s) \right)^{m-1}$$

from which the claim of the lemma follows. ■

A.5 Proof of Theorem 5.2

Proof: We analyze the cost of the last round. Assume $X_1, \ldots, X_n$ are the points of the previous round and $Y_1, \ldots, Y_m$ the points of the last round. Further assume that these points are drawn independently and uniformly from the unit d-cube. By Proposition 3.4 it suffices to prove for all $x \in [0,1]^d$ that the expected number of balls corresponding to a space-filling curve tour through $Y_1, \ldots, Y_m$ and containing x is constant.

Let L be the tour segment starting at Y, where Y is chosen uniformly at random from $Y_1, \ldots, Y_m$. Let B_T be the corresponding ball. If $x \in B_T$ then $|L| > \|x - Y\|$, thus by Lemma 5.1

$$\begin{aligned}
\mathrm{P}\left[x \in B_T\right] &\leq \mathrm{P}\left[|L| > \|x-Y\|\right] \\
&\leq \int_{[0,1]^d} \left[1 - \frac{d}{c_{\psi}^d} \int_{[0,\|x-y\|]} s^{d-1} ds \right]^{m-1} d\lambda^d(y) \\
&\leq \int_{[0,1]^d} \left[1 - \left(\frac{\|x-y\|}{c_{\psi}} \right)^d \right]^{m-1} d\lambda^d(y).
\end{aligned}$$

Let ω_d denote the surface area of the d-sphere and κ_d the volume of the d-ball. Using polar/spherical coordinates we get

$$\begin{aligned}
\mathrm{P}\left[x \in B_T\right] &\leq \int_{\mathbb{R}^d} \left[1 - \left(\frac{x}{c_\psi}\right)^d\right]^{m-1} d\lambda^d(y) \\
&= \omega_{d-1} \int_{[0,\infty)} r^{d-1} \left[1 - \left(\frac{r}{c_\psi}\right)^d\right]^{m-1} d\lambda(r) \\
&= -\omega_{d-1} \frac{c_\psi^d}{m} \left[1 - \left(\frac{r}{c_\psi}\right)^d\right]^m \Bigg|_{r=0}^{\infty} \\
&\leq \kappa_d \frac{c_\psi^d}{m}.
\end{aligned}$$

Thus the expected number of balls containing x is bounded by

$$\kappa_d c_\psi^d,$$

i.e., the volume of a d-ball of radius c_ψ. ■

A.6 Proof of Theorem 5.3

Proof: As in the proof of Theorem 5.2 for uniformly distributed points, it suffices by Proposition 3.4 to bound $\mathrm{P}\left[x \in B_T\right]$ for a point $x \in \mathbb{R}^2$ and B_T the ball around a random tour segment L with starting point Y. Again we bound $\mathrm{P}\left[x \in B_T\right] \leq \mathrm{P}\left[|L| > \|x - Y\|\right]$ using Lemma 5.1.

Without loss of generality we assume that the points are independently and normally distributed centered at the origin and with the identity matrix as covariance matrix. The density function of a point is

$$g(y) = \frac{1}{2\pi} e^{-\frac{1}{2}\|y\|^2}.$$

For a point Y_i with $1 \leq i \leq m$ the probability that it is farther than a distance r_0 from the origin is

$$\mathrm{P}\left[\|Y_i\| > r_0\right] = e^{-\frac{1}{2}r_0^2}.$$

Therefore, the probability that at least one of m points is farther away than r_0 from the origin is bounded from above by $me^{-\frac{1}{2}r_0^2}$. Thus, for

$$r_0 := 4\sqrt{\log m}$$

the probability that one of the points on the tour is farther than r_0 from the origin is bounded from above by

$$me^{-2\log m} = \frac{1}{m}.$$

The points farther away therefore contribute at most $\frac{1}{m}$ to $\mathrm{P}\left[x \in B_T\right]$ and can be ignored in the following.

It remains to bound the probability $\mathrm{P}\left[x \in B_T, \|Y\| < r_0\right]$, i.e., the case that the starting point of the segment is closer than r_0 to the origin. We first bound $\mathrm{P}\left[|L| > \ell_0, \|Y\| < r_0\right]$ for $\ell_0 := \frac{1}{\sqrt{\log m}}$. Using Lemma 5.1 and spherical coordinates we have

$$\mathrm{P}\left[|L| > \ell_0, \|Y\| < r_0\right] \leq \int_{[0,r_0]} re^{-\frac{1}{2}r^2} \left(1 - \frac{2}{c_\psi^2} \int_{[0,\ell_0]} se^{-\frac{1}{2}(r+s)^2} ds\right)^{m-1} dr. \tag{2}$$

For $r < r_0$ we have $(r + \ell_0)^2 = r^2 + 2\ell_0 r + \ell_0^2 \leq r^2 + 8 + \frac{1}{\log m}$, and therefore

$$e^{-\frac{1}{2}(r+\ell_0)^2} \geq e^{-\frac{1}{2}r^2} c_1 \quad \text{for } c_1 = e^{-9}.$$

This yields

$$\begin{aligned}\left(1 - \frac{2}{c_\psi^2}\int_0^{\ell_0} s e^{-\frac{1}{2}(r+s)^2} ds\right)^{m-1} &\leq \left(1 - \frac{2}{c_\psi^2} c_1 e^{-\frac{1}{2}r^2}\int_0^{\ell_0} s ds\right)^{m-1} \\ &= \left(1 - \frac{c_1}{c_\psi^2} e^{-\frac{1}{2}r^2}\ell_0^2\right)^{m-1}.\end{aligned}$$

Inserting this bound into the bound (2) on $\mathrm{P}[|L| > \ell_0, \|Y\| < r_0]$ yields

$$\mathrm{P}[|L| > \ell_0, \|Y\| < r_0] \leq \int_{[0,r_0]} r e^{-\frac{1}{2}r^2}\left(1 - \frac{c_1}{c_\psi^2} e^{-\frac{1}{2}r^2}\ell_0^2\right)^{m-1} dr.$$

We have

$$\frac{c_1}{c_\psi^2} m \ell_0^2 \cdot r e^{-\frac{1}{2}r^2}\left(1 - \frac{c_1}{c_\psi^2} e^{-\frac{1}{2}r^2}\ell_0^2\right)^{m-1} = \frac{d}{dr}\left(1 - \frac{c_1}{c_\psi^2} e^{-\frac{1}{2}r^2}\ell_o^2\right)$$

and therefore

$$\begin{aligned}\mathrm{P}[|L| > \ell_0, \|Y\| < r_0] &\leq \frac{c_\psi^2}{c_1 m \ell_0^2}\left[\left(1 - \frac{c_1}{c_\psi^2} e^{-\frac{1}{2}r^2}\ell_0^2\right)^m\right]_0^{r_0} \\ &\leq \frac{c_\psi^2}{c_1 m \ell_0^2} = \frac{c_\psi^2}{c_1}\frac{\log m}{m}.\end{aligned}$$

Thus there are an expected number of $O(\log m)$ points for which we do not have a bound on $|L|$. We can handle these points by an additional point location data structure like Kirkpatrick's point location hierarchy. This avoids the boundary analysis for the space-filling curve order. The cost induced by the expected $O(\log m)$ points is in $O(n + \log m \log n)$ where the linear part comes from building the point location data structure and the remaining part from the point location queries.

Finally we bound $\mathrm{P}[|L| > \|x - Y\|, |L| < \ell_0, \|Y\| < r_0]$ for $x \in \mathbb{R}^2$. Since Y is distributed spherically symmetric we may assume that x lies on the positive x-axis, i.e., $x = (\chi, 0)$ with $\chi \leq 0$. Furthermore, the probability is 0 if $\chi > l_0 + r_0$, so we can assume

$$0 \leq \chi \leq l_0 + r_0.$$

Let $B(z, r)$ denote the ball of radius r around z. We have so far

$$\begin{aligned}&\mathrm{P}[|L| > \|x - Y\|, |L| < \ell_0, \|Y\| < r_0] \\ &\leq \frac{1}{2\pi}\int_{B(x,l_0)} e^{-\frac{1}{2}\|y\|^2}\left(1 - \frac{2c_1}{c_\psi^2}\|x - y\|^2 e^{-\frac{1}{2}\|y\|^2}\right)^{m-1} d\lambda(y) \\ &= \frac{1}{2\pi}\int_{B(0,l_0)} e^{-\frac{1}{2}\|x-y'\|^2}\left(1 - \frac{2c_1}{c_\psi^2}\|y'\|^2 e^{-\frac{1}{2}\|x-y'\|^2}\right)^{m-1} d\lambda(y'). \qquad (3)\end{aligned}$$

Now,

$$\begin{aligned}\|x-y'\|^2 \geq \qquad\qquad\qquad\qquad & (\chi - l_0)^2 \;=\; \chi^2 + l_0^2 - 2\chi l_0 \\ \geq \chi^2 + l_0^2 - 2(r_0+l_0)l_0 & \\ \geq \chi^2 - 8 &\end{aligned}$$

by the definition of r_0 and l_0, and analogously

$$\begin{aligned}\|x-y'\|^2 \leq (\chi + l_0)^2 \;&=\; \chi^2 + l_0^2 + 2(r_0+l_0)l_0 \\ &\leq \chi^2 + 8 + 2l_0^2.\end{aligned}$$

Inserting these bounds on $\|x-y'\|^2$ in the bound (3) we obtain

$$\begin{aligned}&\mathrm{P}\left[|L| > \|x-Y\|, |L| < \ell_0, \|Y\| < r_0\right] \\ &\leq \frac{1}{2\pi}\int_{B(0,l_0)} e^{8/2}e^{-\frac{1}{2}\chi^2}\left(1 - \frac{2c_1}{c_\psi^2}\|y'\|^2 e^{-9/2}e^{-\frac{1}{2}\chi^2}\right)^{m-1} d\lambda(y') \\ &= e^4 e^{-\frac{1}{2}\chi^2}\int_0^{l_0} r\left(1 - \frac{2c_1}{c_\psi^2}r^2 e^{-9/2}e^{-\frac{1}{2}\chi^2}\right)^{m-1} dr \\ &= -e^{17/2}\frac{c_\psi^2}{2c_1}\frac{1}{m}\left[\left(1 - \frac{2c_1}{c_\psi^2}r^2 e^{-9/2}e^{-\frac{1}{2}\chi^2}\right)^m\right]_{r=0}^{l_0} \\ &\leq e^{17/2}\frac{c_\psi^2}{2c_1}\frac{1}{m}\end{aligned}$$

which proves the theorem.

■

B History of a Round

We here consider the situation where a history is used to locate a point in the current triangulation starting at a simplex of the previous round. We can again bound the cost for such a point location step for a BRIO relative to a randomized insertion order. Let $p'_B(s)$ denote the probability that a given simplex with s stoppers appears in the last round using an arbitrary BRIO. To bound $p'_B(s)$ relative to a randomized insertion order we consider the setting, where points are sampled to rounds as for the biased randomized insertion order, but are inserted in a random order in a round. Overall this corresponds to a randomized insertion order. Let $p'_R(s)$ denote the probability that a given simplex with s stoppers appears in the last round of this construction. To bound the cost of using the history between certain rounds it suffices to bound $p'_B(s)$ relative to $p'_R(s)$. Using the notation from the proof of Lemma 3.1 we have $p'_B(s) \leq \mathrm{P}[V_1] = \mathrm{P}[V_1]\,\mathrm{P}[U_1] \leq 2^{d+1}\mathrm{P}[V_1]\,\mathrm{P}[U_2] \leq 2^{d+1}\mathrm{P}[V_1 \cup U_2] \leq p'_R(s)$.

Remark B.1 For a point set in $\mathbb{R}^d$ it holds that $p'_B(s) \leq 2^{d+1}p'_R(s)$.